



PCT10

## RAW SEQUENCE LISTING

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:25

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

3 <110> APPLICANT: ZENECA LIMITED  
 5 <120> TITLE OF INVENTION: GST SEQUENCES FROM SOYBEAN AND THEIR USE IN THE  
 6 PRODUCTION OF HERBICIDE RESISTANT PLANTS  
 8 <130> FILE REFERENCE: ZENECA CASE PPD50449/WO  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/088,945A  
 C--> 11 <141> CURRENT FILING DATE: 2002-03-21  
 13 <150> PRIOR APPLICATION NUMBER: GB9922346.3  
 14 <151> PRIOR FILING DATE: 1999-09-21  
 16 <160> NUMBER OF SEQ ID NOS: 43  
 18 <170> SOFTWARE: PatentIn Ver. 2.0  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 499  
 22 <212> TYPE: PRT  
 23 <213> ORGANISM: Glycine max  
 25 <400> SEQUENCE: 1  
 26 Met Ser Gln Pro Leu Thr Thr Asn Ser Val Leu Val Glu Glu Ala Ala  
 27 1 5 10 15  
 29 Ala Asp Gly Asp Ser Ser Ala Ala Ala Pro Pro Leu Phe Asp Tyr His  
 30 20 25 30  
 32 Arg Ile Asp Gln Lys Leu Leu Gln Asn Ile Val Tyr Asp Ala Leu Val  
 33 35 40 45  
 35 Trp Ser Thr Leu Asn Cys Leu Leu Val Gly Asp Lys Ser Val Gln Arg  
 36 50 55 60  
 38 Ser Gly Arg Val Pro Gly Val Gly Leu Val His Leu Pro Leu Ser Leu  
 39 65 70 75 80  
 41 Leu Pro Gly Pro Phe Pro Glu Ser His Trp Lys Gln Gly Cys Glu Leu  
 42 85 90 95  
 44 Ala Pro Ile Phe Asn Glu Leu Val Asp Arg Val Ser Leu Asp Gly Lys  
 45 100 105 110  
 47 Phe Leu Gln Glu Ser Leu Ser Arg Thr Lys Asn Ala Asp Glu Phe Thr  
 48 115 120 125  
 50 Ser Arg Leu Leu Asp Ile His Ser Lys Met Leu Gln Ile Asn Lys Lys  
 51 130 135 140  
 53 Glu Asp Ile Arg Met Gly Ile Val Arg Ser Asp Tyr Met Ile Asp Glu  
 54 145 150 155 160  
 56 Lys Thr Lys Ser Leu Gln Ile Glu Met Asn Thr Ile Ser Thr Ser  
 57 165 170 175  
 59 Phe Ala Leu Ile Gly Cys Leu Met Thr Gly Leu His Lys Ser Leu Leu  
 60 180 185 190  
 62 Ser Gln Tyr Gly Lys Phe Leu Gly Leu Asn Ser Asn Arg Val Pro Ala  
 63 195 200 205  
 65 Asn Asn Ala Val Asp Gln Ser Ala Glu Ala Leu Ala Lys Ala Trp Ser  
 66 210 215 220

ENTERED



## RAW SEQUENCE LISTING

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:25

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

```

68 Glu Tyr Asn Asn Pro Arg Ala Ala Ile Leu Val Val Val Gln Val Glu
69 225                230                235                240
71 Glu Arg Asn Met Tyr Glu Gln His Tyr Ile Ser Ala Leu Leu Arg Glu
72                245                250                255
74 Lys His His Ile Arg Ser Ile Arg Lys Thr Leu Thr Glu Ile Asp Gln
75                260                265                270
77 Glu Gly Lys Ile Leu Pro Asp Gly Thr Leu Ser Val Asp Gly Gln Ala
78                275                280                285
80 Ile Ser Val Val Tyr Phe Arg Ala Gly Tyr Thr Pro Lys Asp Tyr Pro
81 290                295                300
83 Ser Glu Ser Glu Trp Arg Ala Arg Leu Leu Met Glu Gln Ser Ser Ala
84 305                310                315                320
86 Ile Lys Cys Pro Thr Ile Ser Tyr His Leu Val Gly Thr Lys Lys Ile
87                325                330                335
89 Gln Gln Glu Leu Ala Lys Pro Gly Val Leu Glu Arg Phe Val Glu Asn
90                340                345                350
92 Lys Asp His Ile Ala Lys Leu Arg Ala Cys Phe Ala Gly Leu Trp Ser
93                355                360                365
95 Leu Glu Asp Ser Asp Ile Val Lys Lys Ala Ile Glu Asn Pro Glu Leu
96 370                375                380
98 Phe Val Met Lys Pro Gln Arg Glu Gly Gly Gly Asn Asn Ile Tyr Gly
99 385                390                395                400
101 Asp Glu Leu Arg Glu Thr Leu Leu Lys Leu Gln Glu Ala Gly Ser Gln
102                405                410                415
104 Glu Asp Ala Ala Tyr Ile Leu Met Gln Arg Ile Phe Pro Ala Thr Ser
105                420                425                430
107 Pro Ala Ile Leu Val Arg Asp Gly Asn Trp Asp Thr Gly His Val Ile
108                435                440                445
110 Ser Glu Ala Gly Ile Phe Gly Thr Tyr Leu Arg Asn Lys Asp Lys Ile
111 450                455                460
113 Ile Ile Asn Asn Glu Ser Gly Tyr Met Val Arg Thr Lys Ile Ser Ser
114 465                470                475                480
116 Ser Tyr Glu Gly Gly Val Leu Pro Gly Phe Gly Val Val Asp Thr Val
117                485                490                495
119 Tyr Leu Thr
123 <210> SEQ ID NO: 2
124 <211> LENGTH: 10
125 <212> TYPE: PRT
126 <213> ORGANISM: Artificial Sequence
128 <220> FEATURE:
129 <223> OTHER INFORMATION: Description of Artificial Sequence:Protein
130 Fragment
132 <400> SEQUENCE: 2
133 Lys Lys Ile Gln Gln Glu Leu Ala Lys Pro
134 1 5 10
137 <210> SEQ ID NO: 3
138 <211> LENGTH: 8
139 <212> TYPE: PRT
140 <213> ORGANISM: Artificial Sequence

```



## RAW SEQUENCE LISTING

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:25

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

```

142 <220> FEATURE:
143 <223> OTHER INFORMATION: Description of Artificial Sequence:Protein
144     Fragment
146 <400> SEQUENCE: 3
147 Cys Phe Ala Gly Leu Trp Ser Leu
148     1             5
151 <210> SEQ ID NO: 4
152 <211> LENGTH: 15
153 <212> TYPE: PRT
154 <213> ORGANISM: Artificial Sequence
156 <220> FEATURE:
157 <223> OTHER INFORMATION: Description of Artificial Sequence:Protein
158     Fragment
160 <400> SEQUENCE: 4
161 Val Met Lys Pro Gln Arg Glu Gly Gly Gly Asn Asn Ile Tyr Gly
162     1             5             10             15
165 <210> SEQ ID NO: 5
166 <211> LENGTH: 11
167 <212> TYPE: PRT
168 <213> ORGANISM: Artificial Sequence
170 <220> FEATURE:
171 <223> OTHER INFORMATION: Description of Artificial Sequence:Protein
172     Fragment
174 <400> SEQUENCE: 5
175 Ala Ala Tyr Ile Leu Met Gln Arg Ile Phe Pro
176     1             5             10
179 <210> SEQ ID NO: 6
180 <211> LENGTH: 1854
181 <212> TYPE: DNA
182 <213> ORGANISM: Glycine max
184 <400> SEQUENCE: 6
185 ggcggccttg tttgttctac ttcctcttac actgggatta gaacaaggcg tttgtgcact 60
186 tctaacaaca ccacctttcc cttcccccaa caacctcaat cactctcttt cgctaaacct 120
187 ctcaaactca tgtctcaacc tttgaccacc aactctgttc ttgttgaaga ggctgctgct 180
188 gatggtgatt cctccgccgc cgcacctccc ctcttcgatt atcatcgtat cgaccaaaaa 240
189 ctgctccaaa acatagttta cgatgctctt gtctggagca ccctcaactg cctccttggt 300
190 ggtgacaaat ctgttcagag atcaggaaga gtccctgggtg tgggcctggt acatctccca 360
191 ctttccttat tacctgggcc atttcctgaa agtcattgga agcaaggggtg cgaattagct 420
192 cctatattta atgaacttgt tgatcgggtg agtttggtat ggaaatttct ccaggaatct 480
193 ctctccagaa ctaagaatgc ggatgaattt acctcaagac ttttagatat tcattctaag 540
194 atgctacaga ttaacaaaaa agaggacata cgcattggaa tagttcgttc agattatatg 600
195 attgatgaga agactaaatc acttttacia atagagatga acactatttc cacttcattt 660
196 gctttgattg gttgtcttat gactggactt cataagagct tactttctca atatggaaaa 720
197 ttccttggac taaattccaa tagggttctt gccataatg ccgttgatca gtctgcagag 780
198 gccttggcta aagcttggag tgagtataac aatcccaggg ctgcaattct ggtcgtgggt 840
199 caggttgaag aaagaaacat gtacgagcag cattatatat ctgcacttct aagagaaaaag 900
200 catcatatta gaagcatacg caaaacgttg accgaaattg atcaggaagg aaaaattctg 960
201 ccagatggaa cactttctgt ggatggacaa gcaatttcag ttgtttactt ccgggctggc 1020
202 tacacgccaa aggactatcc ttcagaatca gaatggagag ctaggctact gatggaacaa 1080

```



## RAW SEQUENCE LISTING

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:25

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

```

203 tcttctgcta tcaaatgccc tacaatatct tatcatttgg ttggcaccaa aaagattcaa 1140
204 caggaacttg caaagcctgg tgttcttgag aggttcggtg aaaacaaaga ccacattgcc 1200
205 aaattgcgtg catgctttgc agggttgttg agtttggaag actcagatat tgttaaaaaa 1260
206 gcaattgaaa atccagagct atttgtgatg aagcctcaaa gagaaggagg aggaaacaat 1320
207 atttatgggtg atgagttgag ggaaaccctc cttaaattac aggaagcagg ttctcaagaa 1380
208 gatgcagcat acatccttat gcagaggata tttcccgcca cttctccagc aattttgggtg 1440
209 cgtgatggta attgggatac gggtcatgtc atttcagaag ctggaatatt tgggtacttat 1500
210 ttaaggaata aggacaagat tatcattaat aacgaaagtg gctatatggt gcgtacaaaa 1560
211 atatcatcat cttatgaagg aggagttttg cctgggtttg gagtggtaga tactgtatac 1620
212 ctaacttgat ggagctaacc cccaagtta tcaaagcaat tcaaaacatt atgtatgggtt 1680
213 tatatatcac cactcaagtc tcctcactcc tgattttctt tgtatggagg cattgctggtt 1740
214 tcttttaatt gttcctatgg gatgggtgtc aattattaac tgtactcaac gacctgtttg 1800
215 attctaacca ataaagattg atgaactgtt ctaacaaaaa aaaaaaaaaa aaaa 1854
217 <210> SEQ ID NO: 7
218 <211> LENGTH: 222
219 <212> TYPE: PRT
220 <213> ORGANISM: Glycine max
222 <400> SEQUENCE: 7
223 Met Ser Ser Ser Gln Glu Glu Val Thr Leu Leu Gly Val Val Gly Ser
224 1 5 10 15
226 Pro Phe Leu His Arg Val Gln Ile Ala Leu Lys Leu Lys Gly Val Glu
227 20 25 30
229 Tyr Lys Tyr Leu Glu Asp Asp Leu Asn Asn Lys Ser Asp Leu Leu Leu
230 35 40 45
232 Lys Tyr Asn Pro Val Tyr Lys Met Ile Pro Val Leu Val His Asn Glu
233 50 55 60
235 Lys Pro Ile Ser Glu Ser Leu Val Ile Val Glu Tyr Ile Asp Asp Thr
236 65 70 75 80
238 Trp Lys Asn Asn Pro Ile Leu Pro Ser Asp Pro Tyr Gln Arg Ala Leu
239 85 90 95
241 Ala Arg Phe Trp Ala Lys Phe Ile Asp Asp Lys Cys Val Val Pro Ala
242 100 105 110
244 Trp Lys Ser Ala Phe Met Thr Asp Glu Lys Glu Lys Glu Lys Ala Lys
245 115 120 125
247 Glu Glu Leu Phe Glu Ala Leu Ser Phe Leu Glu Asn Glu Leu Lys Gly
248 130 135 140
250 Lys Phe Phe Gly Gly Glu Glu Phe Gly Phe Val Asp Ile Ala Ala Val
251 145 150 155 160
253 Leu Ile Pro Ile Ile Gln Glu Ile Ala Gly Leu Gln Leu Phe Thr Ser
254 165 170 175
256 Glu Lys Phe Pro Lys Leu Ser Lys Trp Ser Gln Asp Phe His Asn His
257 180 185 190
259 Pro Val Val Asn Glu Val Met Pro Pro Lys Asp Gln Leu Phe Ala Tyr
260 195 200 205
262 Phe Lys Ala Arg Ala Gln Ser Phe Val Ala Lys Arg Lys Asn
263 210 215 220
266 <210> SEQ ID NO: 8
267 <211> LENGTH: 235
268 <212> TYPE: PRT

```



## RAW SEQUENCE LISTING

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:25

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

```

269 <213> ORGANISM: Glycine max
271 <400> SEQUENCE: 8
272 Met Ala Glu Gln Asp Lys Val Ile Leu His Gly Met Trp Ala Ser Pro
273   1           5           10           15
275 Tyr Ala Lys Arg Val Glu Leu Ala Leu Asn Phe Lys Gly Ile Pro Tyr
276           20           25           30
278 Glu Tyr Val Glu Glu Asp Leu Arg Asn Lys Ser Asp Leu Leu Lys
279           35           40           45
281 Tyr Asn Pro Val His Lys Lys Val Pro Val Leu Val His Asn Gly Lys
282           50           55           60
284 Ala Ile Ala Glu Ser Met Val Ile Leu Glu Tyr Ile Asp Glu Thr Trp
285   65           70           75           80
287 Lys Asp Gly Pro Lys Leu Leu Pro Ser Asp Ser Tyr Lys Arg Ala Gln
288           85           90           95
290 Ala Arg Phe Trp Cys His Phe Ile Gln Asp Gln Leu Met Glu Ser Thr
291           100          105          110
293 Phe Leu Val Val Lys Thr Asp Gly Glu Ala Gln Gln Lys Ala Ile Asp
294           115          120          125
296 His Val Tyr Glu Lys Leu Lys Val Leu Glu Asp Gly Met Lys Thr Tyr
297           130          135          140
299 Leu Gly Glu Gly Asn Ala Ile Ile Ser Gly Val Glu Asn Asn Phe Gly
300  145          150          155          160
302 Ile Leu Asp Ile Val Phe Cys Ala Leu Tyr Gly Ala Tyr Lys Ala His
303           165          170          175
305 Glu Glu Val Ile Gly Leu Lys Phe Ile Val Pro Glu Lys Phe Pro Val
306           180          185          190
308 Leu Phe Ser Trp Leu Met Ala Ile Ala Glu Val Glu Ala Val Lys Ile
309           195          200          205
311 Ala Thr Pro Pro His Glu Lys Thr Val Gly Ile Leu Gln Leu Phe Arg
312           210          215          220
314 Leu Ser Ala Leu Lys Ser Ser Ser Ala Thr Glu
315  225          230          235
318 <210> SEQ ID NO: 9
319 <211> LENGTH: 223
320 <212> TYPE: PRT
321 <213> ORGANISM: Glycine max
323 <400> SEQUENCE: 9
324 Met Ala Glu Val Lys Leu His Gly Phe Trp Tyr Ser Pro Tyr Thr Leu
325   1           5           10           15
327 Arg Val Val Trp Thr Leu Lys Leu Lys Asp Ile Pro Tyr Gln Asn Ile
328           20           25           30
330 Glu Glu Asp Arg Tyr Asn Lys Ser Leu Gln Leu Leu Glu Tyr Asn Pro
331           35           40           45
333 Val Tyr Lys Lys Thr Pro Val Leu Val His Asn Gly Lys Pro Leu Cys
334           50           55           60
336 Glu Ser Met Leu Ile Val Glu Tyr Ile Asp Glu Ile Trp Ala His Asn
337   65           70           75           80
339 Ser Leu Leu Pro Ala Asp Pro Tyr Glu Arg Ala Leu Ala Arg Phe Trp
340           85           90           95

```



RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 07/16/2003  
PATENT APPLICATION:    US/10/088,945A      TIME: 14:50:26

Input Set : A:\50449 sequence listing.txt  
Output Set: N:\CRF4\07162003\J088945A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:17; N Pos. 16,24



## VERIFICATION SUMMARY

DATE: 07/16/2003

PATENT APPLICATION: US/10/088,945A

TIME: 14:50:26

Input Set : A:\50449 sequence listing.txt

Output Set: N:\CRF4\07162003\J088945A.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number  
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:535 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:17  
L:541 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17 after pos.:0